# Annexe 3.3

PCA with specific OF/F/S

## WS

### Best models

|  |  |  |  |
| --- | --- | --- | --- |
| **Ridge Regression** | **Decision Tree** | **Gradient Boosting** | **Random Forest** |
| Alpha: 1.0  Solver: 'saga' | Criterion: 'friedman\_mse'  Max Features: 'sqrt'  Min Samples Split: 3  Splitter: 'best' | Learning Rate: 0.1  Loss: 'huber'  Number of Estimators: 100  Warm Start: False | Criterion: 'friedman\_mse'  Max Features: 'log2'  Min Samples Split: 2  Number of Estimators: 100 |
| **AdaBoost** | **K-Nearest** | **MLP Regressor** | **Elastic Net** |
| Learning Rate: 0.1  Loss: 'exponential'  Number of Estimators: 100 | Algorithm: 'brute'  Leaf Size: 5  Metric: 'cosine'  Number of Neighbors: 10  Weights: 'distance' | Activation: 'tanh'  Hidden Layer Sizes: (100, 100, 100, 100)  Learning Rate: 'constant'  Solver: 'lbfgs' | Copy X: True  Fit Intercept: True  L1 Ratio: 0.25  Positive: False  Precompute: True  Selection: 'random'  Warm Start: False |
| **SGD Regressor** | **Support Vector** | **Bayesian Ridge** | **Kernel Ridge** |
| Learning Rate: 'invscaling'  Loss: 'squared\_error'  Penalty: 'l1'  Warm Start: False | Degree: 1  Gamma: 'scale'  Kernel: 'poly'  Shrinking: True | Alpha 1: 1e-05  Alpha 2: 1e-07  Lambda 1: 1e-07  Lambda 2: 1e-05 | Alpha: 1.0  Coef0: 1.0  Degree: 2  Kernel: 'poly' |
| **Linear Regression** | **RANSAC** | **TheilSen** | **Tensorflow** |
| Copy X: True  Fit Intercept: True  Positive: False | Loss: 'absolute\_error'  Max Trials: 150  Min Samples: 50 | Max Subpopulation: 1000  Number of Subsamples: None |  |

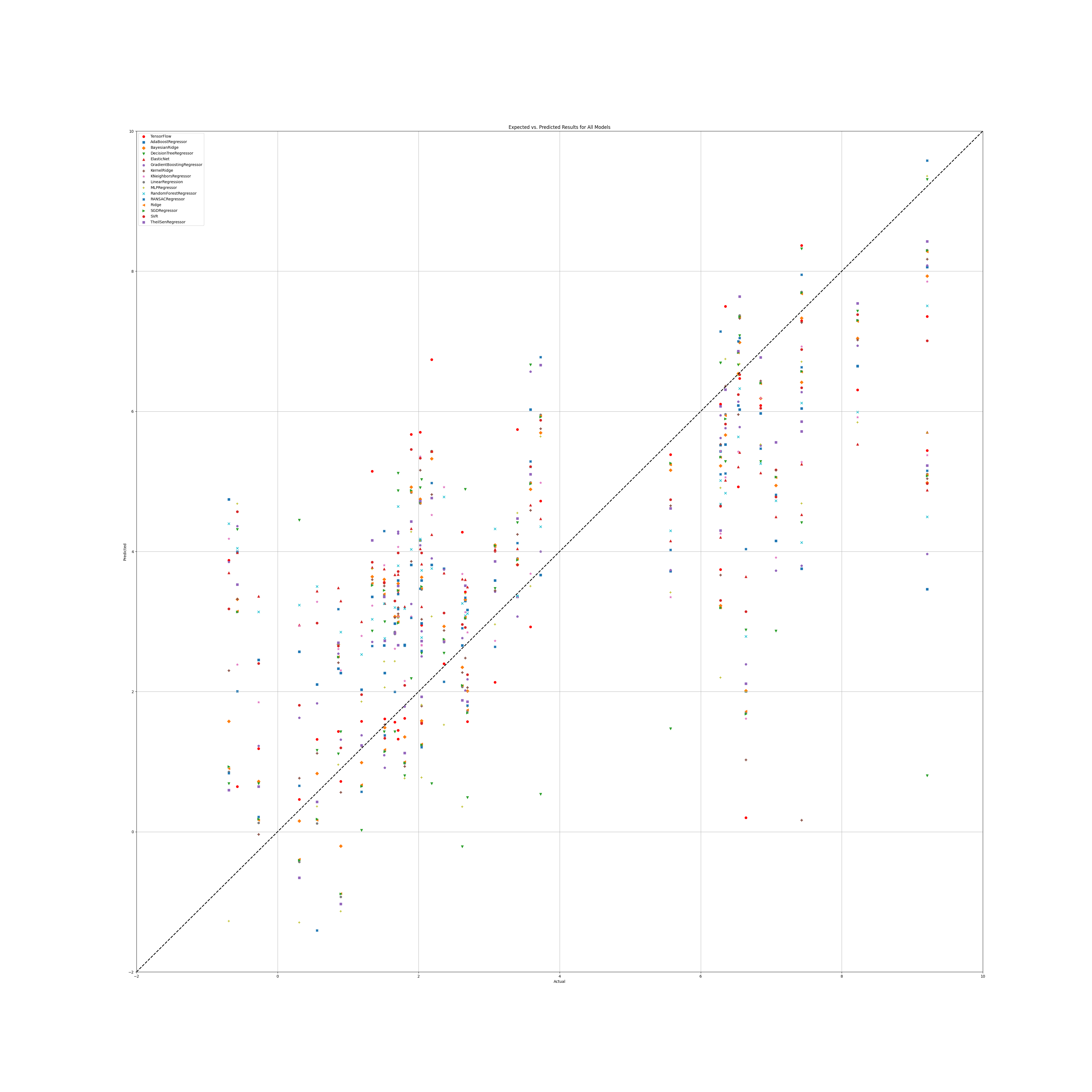
### Prediction Results

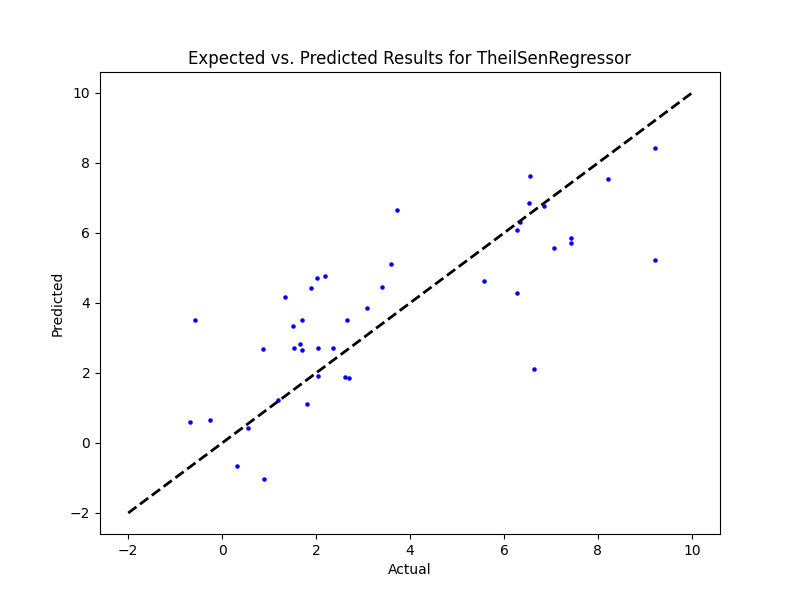
Model: Ridge  
RMSE: 1.8025432740076217  
  
Model: DecisionTreeRegressor  
RMSE: 2.5155145089013833  
  
Model: GradientBoostingRegressor  
RMSE: 2.0313529241787593  
  
Model: RandomForestRegressor  
RMSE: 2.238449339167442  
  
Model: AdaBoostRegressor

RMSE: 2.158289073233721  
  
Model: KNeighborsRegressor  
RMSE: 2.067841525247516  
  
Model: MLPRegressor  
RMSE: 1.9240679844589137

Model: ElasticNet  
RMSE: 2.2939808308547485  
  
Model: SGDRegressor  
RMSE: 1.8081333506047572  
  
Model: SVR  
RMSE: 2.105010334033278  
  
Model: BayesianRidge  
RMSE: 1.8242063487812679  
  
Model: KernelRidge  
RMSE: 2.120151758606026  
  
Model: LinearRegression  
RMSE: 1.8024640289103107  
  
Model: RANSACRegressor  
RMSE: 1.6526001034964783  
  
Model: TheilSenRegressor  
RMSE: 1.7798016464107465  
Model: TensorFlow  
RMSE: 2.097823593885385

### Graphs





## WS Benefit

|  |  |  |  |
| --- | --- | --- | --- |
| **Ridge Regression** | **Decision Tree** | **Gradient Boosting** | **Random Forest** |
|  |  |  |  |
| **AdaBoost** | **K-Nearest** | **MLP Regressor** | **Elastic Net** |
|  |  |  |  |
| **SGD Regressor** | **Support Vector** | **Bayesian Ridge** | **Kernel Ridge** |
|  |  |  |  |
| **Linear Regression** | **RANSAC** | **TheilSen** | **Tensorflow** |
|  |  |  |  |

## Prediction Results

## Graphs

## NR

|  |  |  |  |
| --- | --- | --- | --- |
| **Ridge Regression** | **Decision Tree** | **Gradient Boosting** | **Random Forest** |
| Alpha: 1.0  Solver: 'saga' | Criterion: 'friedman\_mse'  Max Features: 2  Min Samples Split: 5  Splitter: 'random' | Learning Rate: 0.01  Loss: 'absolute\_error'  Number of Estimators: 250  Warm Start: False | Criterion: 'squared\_error'  Max Features: 'log2'  Min Samples Split: 5  Number of Estimators: 50 |
| **AdaBoost** | **K-Nearest** | **MLP Regressor** | **Elastic Net** |
| Learning Rate: 0.001  Loss: 'square'  Number of Estimators: 20 | Algorithm: 'ball\_tree'  Leaf Size: 5  Metric: 'cityblock'  Number of Neighbors: 10  Weights: 'distance' | Activation: 'identity'  Hidden Layer Sizes: (50, 50, 50)  Learning Rate: 'adaptive'  Solver: 'adam' | Copy X: False  Fit Intercept: True  L1 Ratio: 0.25  Positive: True  Precompute: False  Selection: 'random'  Warm Start: False |
| **SGD Regressor** | **Support Vector** | **Bayesian Ridge** | **Kernel Ridge** |
| Learning Rate: 'constant'  Loss: 'squared\_error'  Penalty: 'l2'  Warm Start: True | Degree: 1  Gamma: 'auto'  Kernel: 'sigmoid'  Shrinking: True | Alpha 1: 1e-07  Alpha 2: 1e-05  Lambda 1: 1e-05  Lambda 2: 1e-07 | Alpha: 1.0  Coef0: 1.0  Degree: 1  Kernel: 'poly' |
| **Linear Regression** | **RANSAC** | **TheilSen** | **Tensorflow** |
| Copy X: True  Fit Intercept: True  Positive: True | Loss: 'squared\_error'  Max Trials: 10  Min Samples: 10 | Max Subpopulation: 1000  Number of Subsamples: None |  |

## Prediction results

Model: Ridge  
RMSE: 2.672225538680086  
  
Model: DecisionTreeRegressor  
RMSE: 3.9845306028915664  
  
Model: GradientBoostingRegressor  
RMSE: 2.6163119384479248  
  
Model: RandomForestRegressor  
RMSE: 2.6470511051432  
  
Model: AdaBoostRegressor  
RMSE: 2.688978656767679  
  
Model: KNeighborsRegressor  
RMSE: 2.68314426153918  
  
Model: MLPRegressor  
RMSE: 2.67912664737191  
  
Model: ElasticNet  
RMSE: 2.9792360189276654  
  
Model: SGDRegressor  
RMSE: 2.5602691545984078  
  
Model: SVR  
RMSE: 2.843246085988345  
  
Model: BayesianRidge  
RMSE: 2.6966714264564704  
  
Model: KernelRidge  
RMSE: 2.6658793134028334  
  
Model: LinearRegression  
RMSE: 2.7968966479856294  
  
Model: RANSACRegressor  
RMSE: 4.126794940669215  
  
Model: TheilSenRegressor  
RMSE: 3.2156799342479787  
  
Model: TensorFlow  
RMSE: 3.6905011881976604

## Graph

## NR Benefit

|  |  |  |  |
| --- | --- | --- | --- |
| **Ridge Regression** | **Decision Tree** | **Gradient Boosting** | **Random Forest** |
|  |  |  |  |
| **AdaBoost** | **K-Nearest** | **MLP Regressor** | **Elastic Net** |
|  |  |  |  |
| **SGD Regressor** | **Support Vector** | **Bayesian Ridge** | **Kernel Ridge** |
|  |  |  |  |
| **Linear Regression** | **RANSAC** | **TheilSen** | **Tensorflow** |
|  |  |  |  |

## Prediction results

## Graph

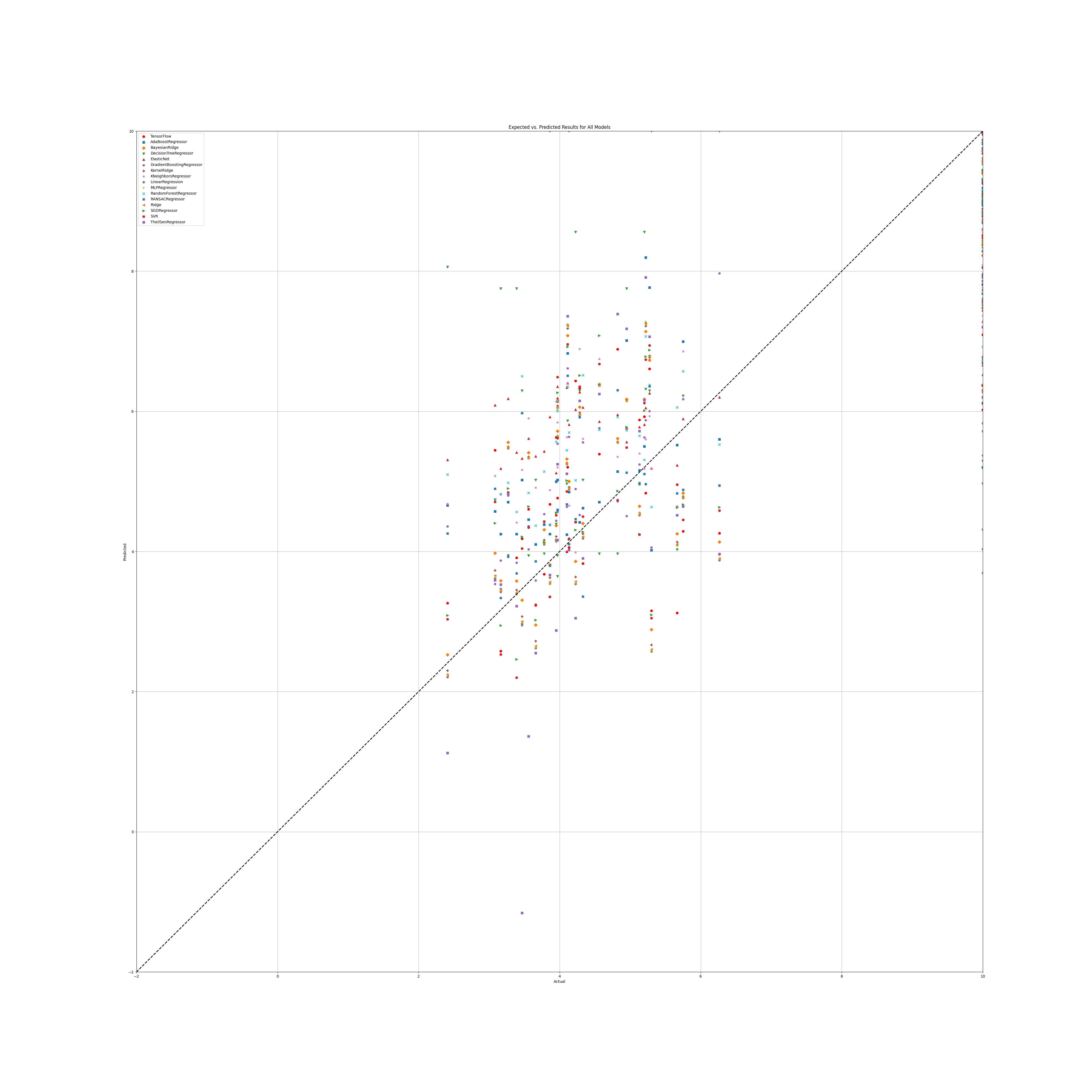
## PR

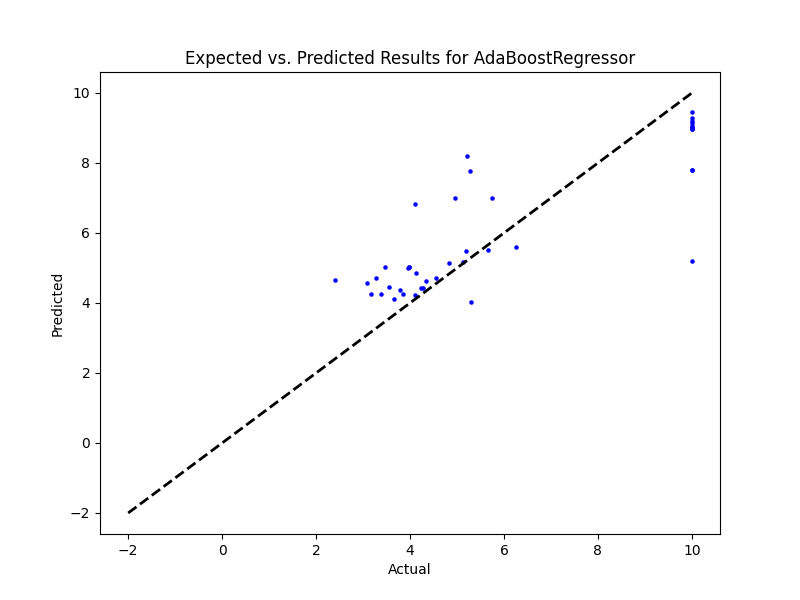
|  |  |  |  |
| --- | --- | --- | --- |
| **Ridge Regression** | **Decision Tree** | **Gradient Boosting** | **Random Forest** |
| Alpha: 1.0  Solver: 'sparse\_cg' | Criterion: 'friedman\_mse'  Max Features: 'log2'  Min Samples Split: 5  Splitter: 'random' | Learning Rate: 0.1  Loss: 'huber'  Number of Estimators: 100  Warm Start: True | Criterion: 'poisson'  Max Features: 'sqrt'  Min Samples Split: 2  Number of Estimators: 50 |
| **AdaBoost** | **K-Nearest** | **MLP Regressor** | **Elastic Net** |
| Learning Rate: 1.0  Loss: 'square'  Number of Estimators: 100 | Algorithm: 'brute'  Leaf Size: 5  Metric: 'cosine'  Number of Neighbors: 10  Weights: 'distance' | Activation: 'identity'  Hidden Layer Sizes: (100, 100, 100, 100)  Learning Rate: 'constant'  Solver: 'sgd' | Copy X: False  Fit Intercept: True  L1 Ratio: 0.25  Positive: False  Precompute: True  Selection: 'random'  Warm Start: False |
| **SGD Regressor** | **Support Vector** | **Bayesian Ridge** | **Kernel Ridge** |
| Learning Rate: 'constant'  Loss: 'epsilon\_insensitive'  Penalty: None  Warm Start: True | Degree: 1  Gamma: 'scale'  Kernel: 'linear'  Shrinking: False | Alpha 1: 1e-05  Alpha 2: 1e-07  Lambda 1: 1e-07  Lambda 2: 1e-05 | Alpha: 0.1  Coef0: 1.0  Degree: 1  Kernel: 'poly' |
| **Linear Regression** | **RANSAC** | **TheilSen** | **Tensorflow** |
| Copy X: True  Fit Intercept: True  Positive: False | Loss: 'absolute\_error'  Max Trials: 50  Min Samples: 50 | Max Subpopulation: 1000  Number of Subsamples: None |  |

## Prediction Results

Model: Ridge  
RMSE: 1.5428936128793769  
  
Model: DecisionTreeRegressor  
RMSE: 3.2469196463627448  
  
Model: GradientBoostingRegressor  
RMSE: 1.4367618941263114  
  
Model: RandomForestRegressor  
RMSE: 1.7401753023379782  
  
Model: AdaBoostRegressor  
RMSE: 1.4601792454403437  
  
Model: KNeighborsRegressor  
RMSE: 1.569952620679771  
  
Model: MLPRegressor  
RMSE: 1.5498178729563896  
  
Model: ElasticNet  
RMSE: 2.020148194117373  
  
Model: SGDRegressor  
RMSE: 1.3982044258473618  
  
Model: SVR  
RMSE: 1.3792349072069061  
  
Model: BayesianRidge  
RMSE: 1.500299816449045  
  
Model: KernelRidge  
RMSE: 1.5296078829409718  
  
Model: LinearRegression  
RMSE: 1.549543337765582  
  
Model: RANSACRegressor  
RMSE: 1172.597357450557  
  
Model: TheilSenRegressor  
RMSE: 2.317243574042818  
  
Model: TensorFlow  
RMSE: 2.135240128991088

## Graph





## PR Benefit

|  |  |  |  |
| --- | --- | --- | --- |
| **Ridge Regression** | **Decision Tree** | **Gradient Boosting** | **Random Forest** |
|  |  |  |  |
| **AdaBoost** | **K-Nearest** | **MLP Regressor** | **Elastic Net** |
|  |  |  |  |
| **SGD Regressor** | **Support Vector** | **Bayesian Ridge** | **Kernel Ridge** |
|  |  |  |  |
| **Linear Regression** | **RANSAC** | **TheilSen** | **Tensorflow** |
|  |  |  |  |

## Prediction Results

## Graphs

## SR

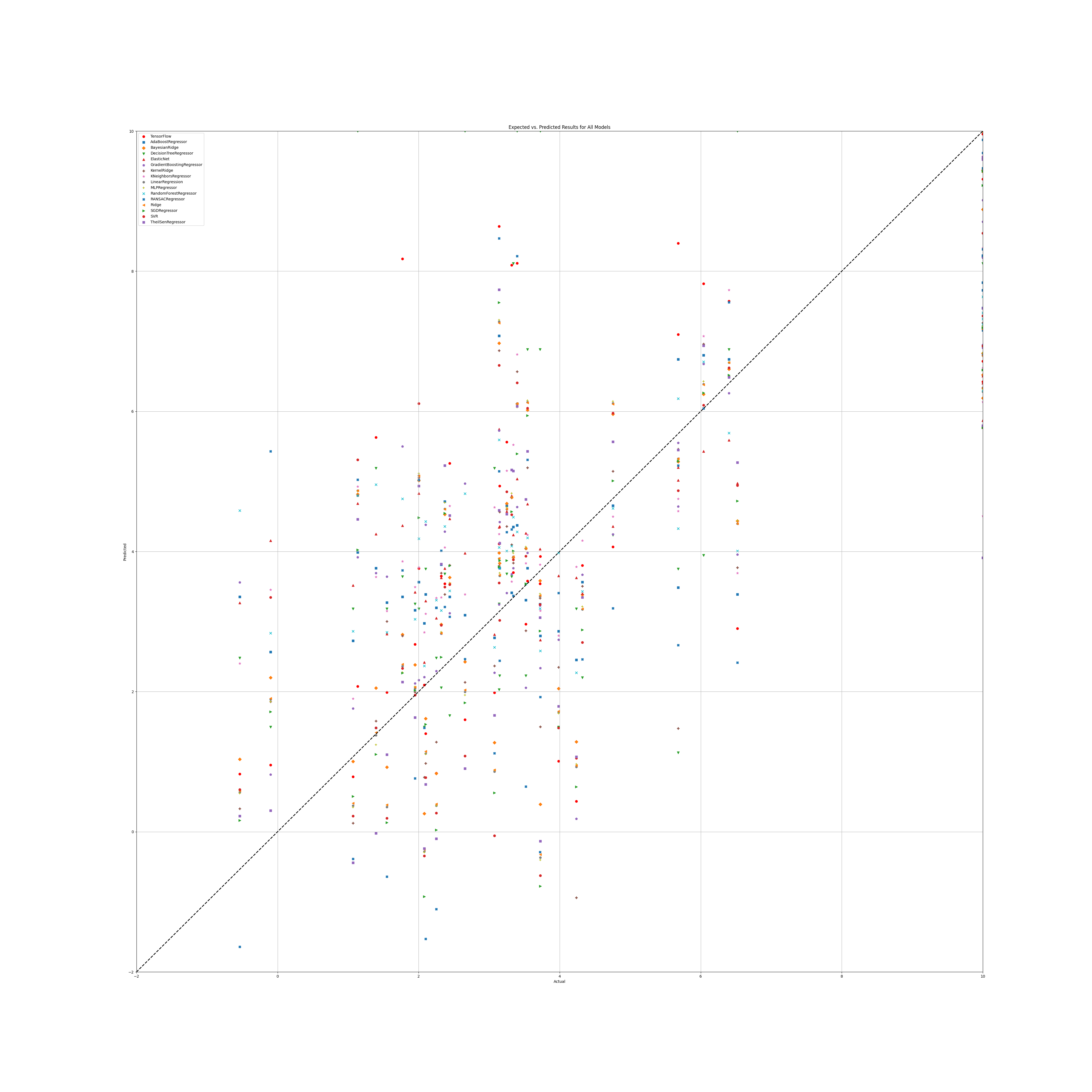
|  |  |  |  |
| --- | --- | --- | --- |
| **Ridge Regression** | **Decision Tree** | **Gradient Boosting** | **Random Forest** |
| Alpha: 1.0  Solver: 'saga' | Criterion: 'squared\_error'  Max Features: 'sqrt'  Min Samples Split: 5  Splitter: 'random' | Learning Rate: 0.1  Loss: 'squared\_error'  Number of Estimators: 250  Warm Start: False | Criterion: 'squared\_error'  Max Features: 'sqrt'  Min Samples Split: 5  Number of Estimators: 50 |
| **AdaBoost** | **K-Nearest** | **MLP Regressor** | **Elastic Net** |
| Learning Rate: 1.0  Loss: 'exponential'  Number of Estimators: 50 | Algorithm: 'brute'  Leaf Size: 5  Metric: 'cosine'  Number of Neighbors: 10  Weights: 'uniform' | Activation: 'identity'  Hidden Layer Sizes: (50, 50, 50)  Learning Rate: 'adaptive'  Solver: 'adam' | Copy X: True  Fit Intercept: True  L1 Ratio: 0.25  Positive: False  Precompute: True  Selection: 'random'  Warm Start: False |
| **SGD Regressor** | **Support Vector** | **Bayesian Ridge** | **Kernel Ridge** |
| Learning Rate: 'constant'  Loss: 'squared\_error'  Penalty: 'elasticnet'  Warm Start: True | Degree: 1  Gamma: 'scale'  Kernel: 'linear'  Shrinking: False | Alpha 1: 1e-05  Alpha 2: 1e-07  Lambda 1: 1e-07  Lambda 2: 1e-05 | Alpha: 1.0  Coef0: 1.0  Degree: 2  Kernel: 'poly' |
| **Linear Regression** | **RANSAC** | **TheilSen** | **Tensorflow** |
| Copy X: True  Fit Intercept: True  Positive: False | Loss: 'absolute\_error'  Max Trials: 50  Min Samples: 50 | Max Subpopulation: 1000  Number of Subsamples: None |  |

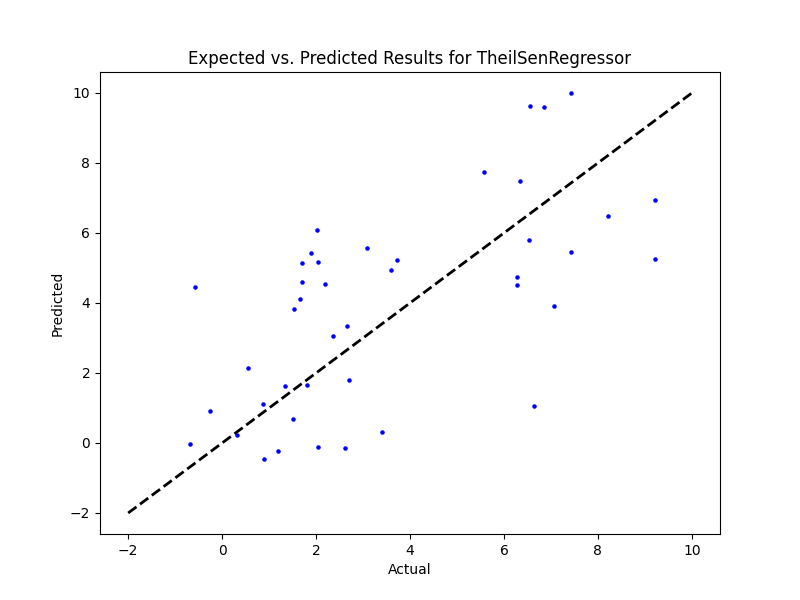
## Prediction Results

Model: Ridge  
RMSE: 2.18139925027502

Model: DecisionTreeRegressor  
RMSE: 2.9678517796866077  
  
Model: GradientBoostingRegressor  
RMSE: 1.767670613841336  
  
Model: RandomForestRegressor  
RMSE: 1.9832456728396721  
  
Model: AdaBoostRegressor  
RMSE: 1.730223643399416  
  
Model: KNeighborsRegressor  
RMSE: 1.9881795317796804  
  
Model: MLPRegressor  
RMSE: 2.1954197419578194  
  
Model: ElasticNet  
RMSE: 2.143238817396353  
  
Model: SGDRegressor  
RMSE: 2.1775358331876498  
  
Model: SVR  
RMSE: 2.26315199305162  
  
Model: BayesianRidge  
RMSE: 2.0402747052702432  
  
Model: KernelRidge  
RMSE: 1.9897716768737832  
  
Model: LinearRegression  
RMSE: 2.19248477142144  
  
Model: RANSACRegressor  
RMSE: 2.5635208359186454  
  
Model: TheilSenRegressor  
RMSE: 3.3807974052830247  
  
Model: TensorFlow  
RMSE: 2.4707587220751748

## Graph





## SR Benefit

|  |  |  |  |
| --- | --- | --- | --- |
| **Ridge Regression** | **Decision Tree** | **Gradient Boosting** | **Random Forest** |
|  |  |  |  |
| **AdaBoost** | **K-Nearest** | **MLP Regressor** | **Elastic Net** |
|  |  |  |  |
| **SGD Regressor** | **Support Vector** | **Bayesian Ridge** | **Kernel Ridge** |
|  |  |  |  |
| **Linear Regression** | **RANSAC** | **TheilSen** | **Tensorflow** |
|  |  |  |  |

## Prediction Results

## Graph